Policy Control for Connected and Tethered Devices

A Sandvine Technology Showcase

Contents

Executive Summary

In this era of the Internet of Things, there is no practical limit to the number of devices that can have an IP address. In addition to 'traditional' devices, the Internet will soon see the addition of billions of new devices as connected homes and other machine-to-machine devices come online.

The increasing number and diversity of connected devices brings opportunity to communications service providers (CSPs) who can identify trends and can create services that anticipate new needs.

To take full advantage of these opportunities, however, and to enable the Internet of Things, the policy control solutions on which CSPs rely must meet a range of technical requirements.

For instance, the policy control platform must be able to identify and distinguish client devices in real-time, and must do so for devices behind network address translators (NATs), devices that are tethered onto the mobile network.

Sandvine's traffic classification technology overcomes all of these challenges to provide CSPs with the insight to understand the rich assortment of devices on their network and to empower CSPs with the policy control capabilities to profitably manage the Internet of Things.

Introduction to Device Awareness

In this era of the Internet of Things, there is no practical limit to the number of devices that can have an IP address. In addition to 'traditional' devices like laptops, tablets, mobile phones, gaming consoles, smart televisions, etc., the Internet will soon see the addition of billions of new devices as connected homes (e.g., thermostats, entertainment devices, security cameras, small appliances, etc.) and other machine-to-machine devices (e.g., fleet management, bank machines, vending machines, smart meters, traffic lights, drones, connected cars, etc.) come online.

The increasing number and diversity of connected devices brings opportunity to communications service providers (CSPs) who can identify trends and can create services that anticipate new needs.

To take full advantage of these opportunities, and to enable the Internet of Things, CSPs must deploy solutions that meet a range of technical requirements, including:[1]

- The ability to identify individual client devices in real-time
- The ability to provide detailed measurements and metrics per device
- The ability to differentiate between a *client device* (the device that originates packets on the network) and an *access device* (the devices that connects to the access network and owns the IP connectivity session)
- The ability to detect device tethering (e.g., when a mobile phone connects to the mobile network and serves as a hotspot for other devices), and to apply separate policy management to the tethered device versus the access device
- The ability to identify devices that are behind equipment performing network address translation (e.g., a home router or public WiFi access point)
- A policy engine that links all measured conditions to real-time policy control (i.e., decision-making and policy enforcement)

Only by meeting all of these requirements can a policy control solution truly give CSPs the business intelligence insight and policy enforcement versatility needed to understand and to meet the growing needs of this evolving connected world.

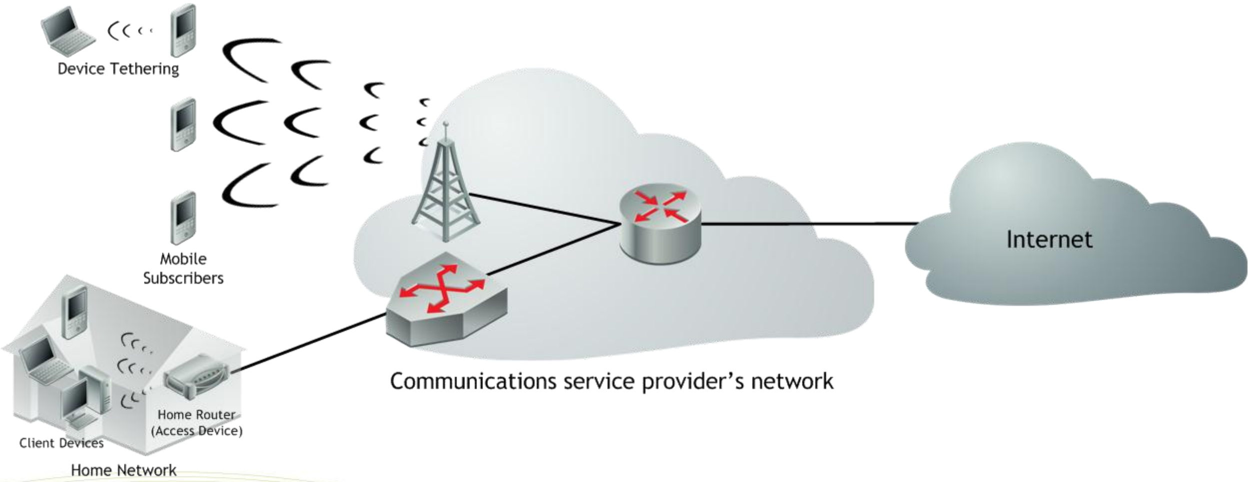

Device Tethering
Mobile Subscribers
Client Devices
Home Router (Access Device)
Home Network
Communications service provider's network
Internet

Figure 1 - 'Traditional' connected devices; notice the device tethering and home-roaming

[1] These requirements are discussed in greater detail in the whitepaper *Internet Traffic Classification*, available at www.sandvine.com

Sandvine's Device Awareness

The Sandvine Policy Engine delivers, and subsequently takes advantage of, real-time device awareness in any network by meeting all of the technical requirements listed previously. The following subsections explain how our solution meets each need.

To take advantage of this unique visibility, Sandvine's customers need only install the Network Business Intelligence Policy Pack.[2]

Fixed Access Networks

Within the home network depicted in Figure 1, there are many client devices (e.g., laptop, tablet, mobile phone), and the diagram could have included many others (e.g., game console, smart thermostat, etc.), but there is only a single access device (i.e., home router). The home router connects to the CSP's network, but the client devices actually originate packets.

At the highest level, client device awareness in fixed access networks is achieved by tracking known patterns within TCP traffic across multiple flows within a subscriber session. Device statistics are counted per-device, per-subscriber.

To recognize client devices, we rely on layer 4 (i.e., transport) attributes to classify the device instance, then examine the User Agent to identify additional features of the device. The process of client device recognition is made up of two main traffic characteristics:

- **The TCP Timestamp Option (RFC 1323[3]):** Many operating systems use this TCP option. It can be used to classify each new TCP flow to a known group of TCP flows that represent a single device.
- **The TCP Source Port:** Some operating systems (e.g., Windows) increment the source port for each new TCP port. This behavior can be used to classify each new TCP flow to a known group of TCP flows that represent a single device.

Additional SandScript[4] policy is applied at the application layer to correlate usage metrics to a *device instance*. Once these initial groupings are defined, a device instance is created and tracked by our Policy Engine. This device instance is identified as a specific type of device by analyzing HTTP User-Agent headers.

This level of granular visibility opens up many possibilities for innovative CSPs; two popular use cases for device awareness in fixed access networks are:

- Third-party/partner promotions: for instance, a CSP can partner with a consumer electronics brand to enable sponsored data connectivity such that the traffic associated with a particular device (or brand of devices) does not count against a subscriber's monthly quota[5]
- Business intelligence insight to understand the growing number of connected devices within a home, including machine-to-machine and home roaming[6] (i.e., using mobile devices on the home WiFi network)

[2] This policy pack is a free install for any Sandvine customer – just grab it and install as you would any software update.
[3] *"TCP Extensions for High Performance"*, available at https://www.ietf.org/rfc/rfc1323.txt
[4] SandScript is Sandvine's event-driven policy definition language, used to turn business rules into network policy control instructions; more information is available at https://www.sandvine.com/technology/sandscript.html
[5] Does this sound farfetched? Vox Telecom in South Africa has already done it, via a partnership with Samsung and using the Sandvine platform

Mobile Access Networks

In the mobile network depicted in Figure 1, things become a bit blurred. Typically, any device that connects to the mobile network is an access device and, in most cases, that same mobile device is also a client device. However, in the case of tethering, a clear split is made: in this case, the mobile phone serves as an access device (as a WiFi hotspot), while the tethered laptop is the client device.

Forgetting tethering for a moment, let's just consider the case in which the mobile device serves as both client and access device.

In 3GPP mobile networks (UMTS, HSPA or LTE), our solution inspects the device's International Mobile Station Equipment Identity (IMEI) code[7], which contains a Type Allocation Code (TAC)[8].

In 3GPP2 (e.g., CDMA2000) networks, we retrieve the Mobile Equipment Identifier (MEID), which resolves to the manufacturer level.

By leveraging a licensed database of devices registered with the GSM Association (GSMA), we are then able to look up the retrieved code to obtain device identification and characteristics.

This method has the added benefit that, in addition to device identification, we gain and leverage knowledge of device characteristics including operation system, radio capability, manufacturer, screen resolution, screen size, etc.

Mobile Tethering

Many CSPs want to offer tethering services as add-ons to existing data plans, but to do so they need to be able to detect and manage tethered devices. The most robust plans require policy control platforms that can apply separate policy to the tethered and access (i.e., hotspot) devices.

Accurate detection of tethered PC devices is a challenge. The current generation of network access nodes (e.g., GGSN for GSM/3G and PDSN for CDMA) don't provide advanced inspection capabilities and can only inspect the HTTP user agent field. The HTTP user agent field, while useful in many contexts, is not sufficient in itself to allow detection of tethering for all tethered device data traffic flow cases. Additionally, the HTTP user agent field can easily be spoofed by the tethered devices to get around any gateway-based policy. The detection of PC device tethering is a non-trivial problem, and requires measurement of many different factors for accurate identification – mistakes ruin customer experience and increase support costs.

The unique freeform design of SandScript allows real-time analysis and stateful tracking of HTTP flows and layer 7 applications to match a specific 'tethering profile'. The Sandvine Policy Engine is powerful enough to interact with multiple real-time events and to invoke policy rules that precisely define device classification by correlating events and measurements for both TCP and UDP flows. As a result, Sandvine's tethering detection is based on matching behavior against predetermined tethering behavioral profiles, based on measurements and observations of, among other things:

- number of simultaneous sessions

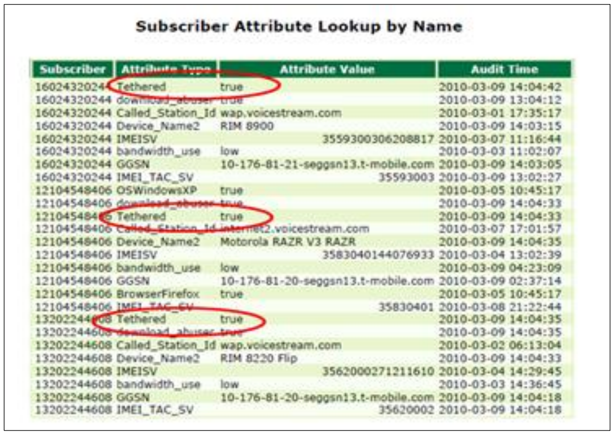
[6] As reported in our 1H2012 Global Internet Phenomena Report, even back in 2012 mobile devices already accounted for almost ten percent of fixed access Internet usage in North America. This phenomenon is driven largely by video and audio streaming: more screens and speakers = more data usage!

[7] You can learn more about IMEI here: http://en.wikipedia.org/wiki/International_Mobile_Station_Equipment_Identity

[8] You can learn more about the Type Allocation Code, and can see some examples, here: http://en.wikipedia.org/wiki/Type_Allocation_Code

- HTTP user-agent headers
- device type
- device screen size
- TCP Timestamp
- TCP Source Port
- TCP Sequence Number
- Application-based correlation to TCP flows
- Node-pair correlation to TCP flows

By correlating all of these measurements, we are able to confidently classify client devices when they are tethered to a separate access device, even when the traffic itself is encrypted.



Subscriber Attribute Lookup by Name

Subscriber	Attribute Type	Attribute Value	Audit Time
16024320244	Tethered	true	2010-03-09 14:04:42
16024320244	download_abuser	true	2010-03-09 13:04:12
16024320244	Called_Station_Id	wap.voicestream.com	2010-03-01 17:35:17
16024320244	Device_Name2	RIM 8900	2010-03-09 14:03:15
16024320244	IMEISV	3559300306208817	2010-03-07 11:16:44
16024320244	bandwidth_use	low	2010-03-03 11:02:07
16024320244	GGSN	10-176-81-21-seggsn13.t-mobile.com	2010-03-09 14:03:05
16024320244	IMEI_TAC_SV	35593003	2010-03-09 13:02:27
12104548406	OSWindowsXP	true	2010-03-05 10:45:17
12104548406	download_abuser	true	2010-03-09 14:04:33
12104548406	Tethered	true	2010-03-09 14:04:33
12104548406	Called_Station_Id	internet2.voicestream.com	2010-03-07 17:01:57
12104548406	Device_Name2	Motorola RAZR V3 RAZR	2010-03-09 14:04:35
12104548406	IMEISV	3583040144076933	2010-03-04 13:02:39
12104548406	bandwidth_use	low	2010-03-09 04:23:09
12104548406	GGSN	10-176-81-20-seggsn13.t-mobile.com	2010-03-09 02:37:14
12104548406	BrowserFirefox	true	2010-03-05 10:45:17
12104548406	IMEI_TAC_SV	35830401	2010-03-08 21:22:44
13202244608	Tethered	true	2010-03-09 14:04:35
13202244608	download_abuser	true	2010-03-09 14:04:35
13202244608	Called_Station_Id	wap.voicestream.com	2010-03-02 06:13:04
13202244608	Device_Name2	RIM 8220 Flip	2010-03-09 14:04:33
13202244608	IMEISV	3562000271211610	2010-03-04 14:29:45
13202244608	bandwidth_use	low	2010-03-03 14:36:45
13202244608	GGSN	10-176-81-20-seggsn13.t-mobile.com	2010-03-09 14:04:18
13202244608	IMEI_TAC_SV	35620002	2010-03-09 14:04:18

Figure 2 - Network Demographics report showing characteristics of tethered devices

Network Address Translators

Network address translators remap one range of IP addresses to another. Consider three examples:

- A home WiFi router (access device) has a single public IP address (provided by the CSP) on the Internet side, and any number of client devices (each with its own internal network IP address) within the home
- A public WiFi access point performs in much the same manner as the home WiFi router
- A mobile phone (access device) that is configured to run as a hotspot to allow a tablet (client device) to use the mobile network; the mobile phone has a public IP address provided by the CSP, while the tablet (and any other devices that connect to the hotspot) have private addressing

To any network equipment on the public side of the NAT, all the devices on the private side share a single IP address. As a consequence, device awareness in the context of a NAT has two distinct components:

1. Distinguishing between the client devices that are concurrently using the NAT at any point in time
2. Providing an identity for each of those client devices

Distinguishing between Concurrent Devices behind a NAT

To determine how many devices are behind a NAT, the Sandvine system relies on the behavior of TCP. Specifically, when a TCP stack initiates on a device (e.g., when the device is turned on and the process begins), a timer starts at a random value.[9] Each time a packet is sent, it is stamped with the current value of the timer.

Figure 3 shows how TCP timestamps can be used to determine how many devices are currently sending traffic from behind a NAT. In this case, Device 1’s TCP stack started at (time = 4) with (value = 181); Device 2’s TCP stack started at (time = 13) with (value = 706), and so on.

If we were to inspect traffic at each point in time and plot the observed TCP timestamps, we would ultimately get four separate curves/lines, each of which corresponds to a distinct device.[10]

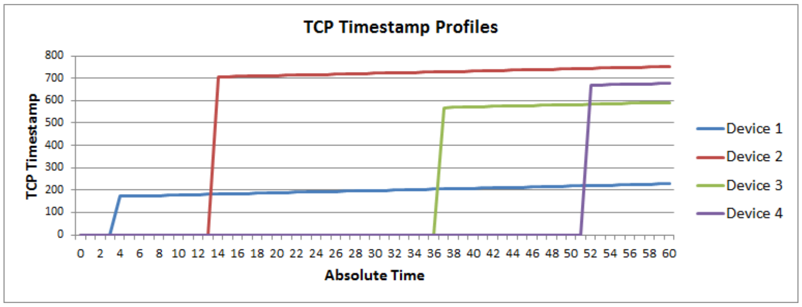
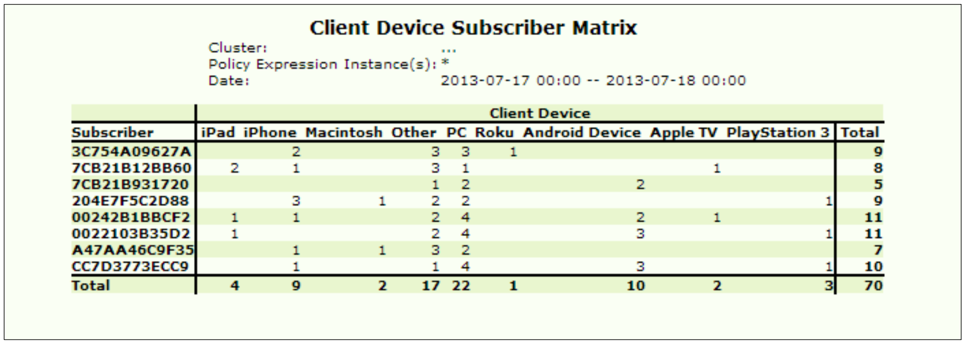


Figure 3 - Four distinct devices result in four distinct timestamp curves/lines.

Identifying Individual Devices behind a NAT

From examining TCP timestamps, the Sandvine system can determine how many devices are concurrently sending traffic through a NAT; the next step is to identify each of these devices.

To do so, the system relies on a range of factors, in a manner very similar to that applied to identifying client devices used in tethering. These behavioral profiles are then compared against a known library, and each profile is updated to reflect the determined identity (e.g., Samsung Smart TV, iPhone 6, PS4, etc.).



Client Device Subscriber Matrix

Cluster: ...
Policy Expression Instance(s): *
Date: 2013-07-17 00:00 -- 2013-07-18 00:00

	Client Device									
Subscriber	iPad	iPhone	Macintosh	Other	PC	Roku	Android Device	Apple TV	PlayStation 3	Total
3C754A09627A		2		3	3	1				9
7CB21B12BB60	2	1		3	1			1		8
7CB21B931720				1	2		2			5
204E7F5C2D88		3	1	2	2				1	9
00242B1BBCF2	1	1		2	4		2	1		11
0022103B35D2	1			2	4		3		1	11
A47AA46C9F35		1	1	3	2					7
CC7D3773ECC9		1		1	4		3		1	10
Total	4	9	2	17	22	1	10	2	3	70

Figure 4 - Network Demographics report showing the number and types of client devices within each subscriber household

[9] You can learn more about TCP timestamps at http://en.wikipedia.org/wiki/Transmission_Control_Protocol#TCP_timestamps

[10] Since the TCP timestamp is a four byte value, it is extremely unlikely that any two devices would choose random TCP stack times that are in alignment (which would result in overlapping profiles). It is worthwhile to note, however, that TCP timestamps do not apply to UDP traffic, so any devices that use UDP *exclusively* will not be detected by this method.

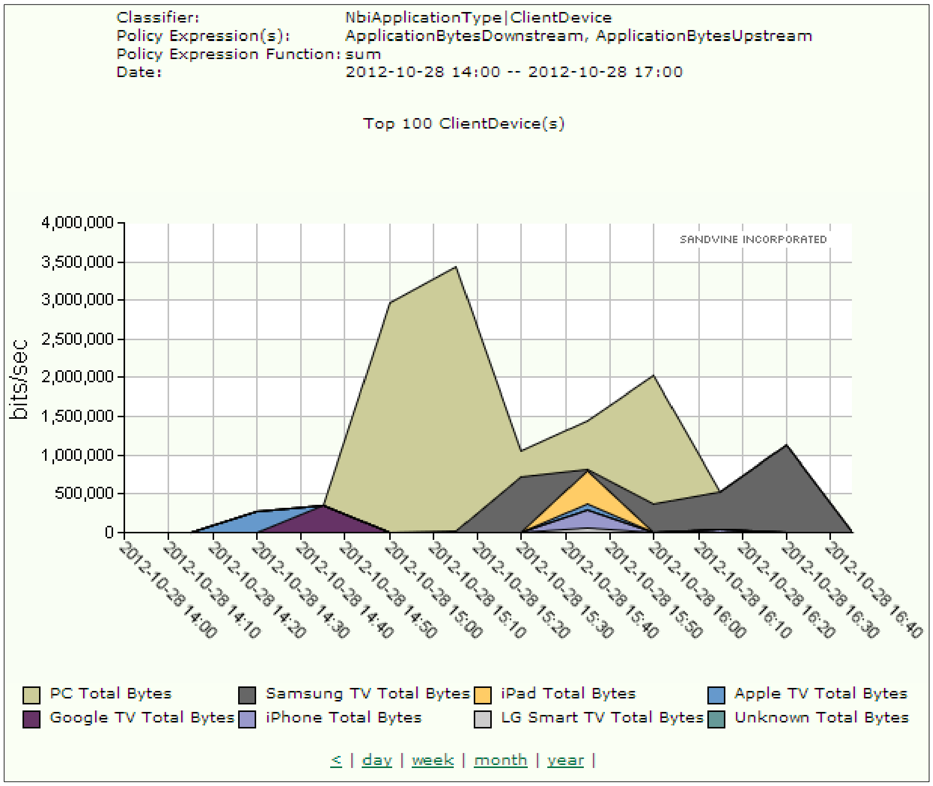
Policy Control and Business Intelligence

The Sandvine Policy Engine not only detects and identifies these device characteristics, but also makes them available as conditions for real-time network policy control. This means that CSPs can make decisions based on and apply management to individual devices, classes of device, etc. to enable an infinite range of use cases.

Furthermore, the measurements and metrics produced by the Sandvine Policy Engine provide vital business intelligence that can be accessed through a variety of interfaces:

- Use our Network Demographics reporting interface to examine historical usage[11]
- Detect and project trends by using our Network Analytics product and its built-in Device Insights Dashboard[12]
- Extract data directly from the Sandvine database using our web services API
- Create configurable-format data records to feed your big data systems[13]

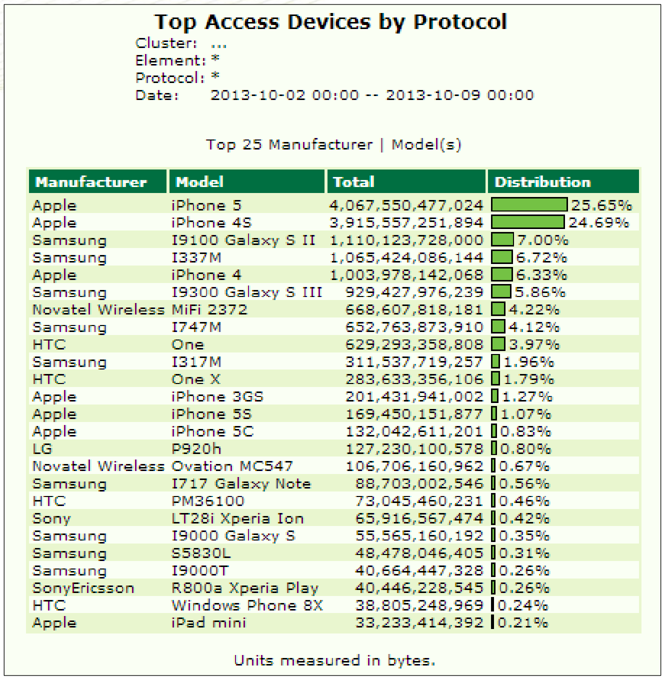
The following images show examples of some of the measurements and metrics available, but are by no means exhaustive. Essentially any Sandvine measurement is available per-device; for a more comprehensive examination of such measurements, please refer to our product documentation or the technology showcase *Internet Traffic Classification* (available at www.sandvine.com).



Classifier: NbiApplicationType|ClientDevice
Policy Expression(s): ApplicationBytesDownstream, ApplicationBytesUpstream
Policy Expression Function: sum
Date: 2012-10-28 14:00 -- 2012-10-28 17:00

Figure 5 - Network Demographics report showing bytes attributable to different device types and manufacturers

[11] More information is available at https://www.sandvine.com/products/network-demographics/
[12] More information is available at https://www.sandvine.com/products/network-analytics/
[13] More information is available at https://www.sandvine.com/solutions/business-intelligence/data-records.html



Top Access Devices by Protocol

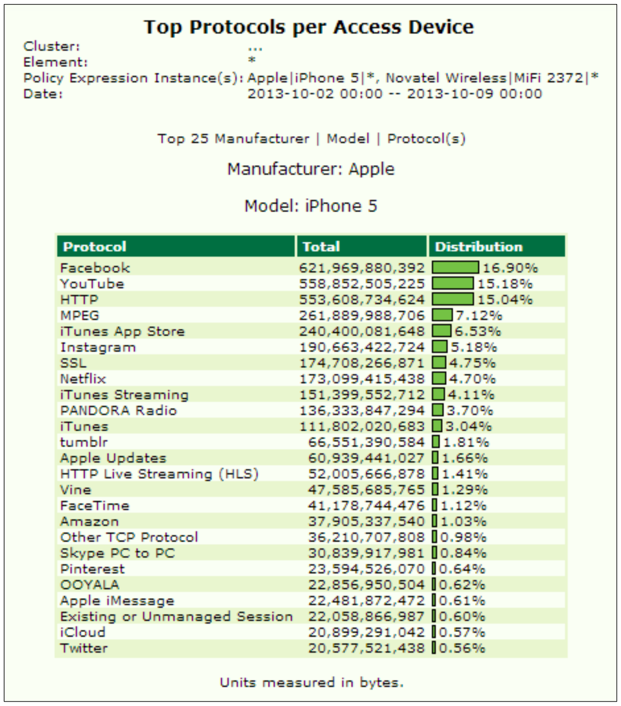
Cluster: ...
Element: *
Protocol: *
Date: 2013-10-02 00:00 -- 2013-10-09 00:00

Top 25 Manufacturer | Model(s)

Manufacturer	Model	Total	Distribution
Apple	iPhone 5	4,067,550,477,024	25.65%
Apple	iPhone 4S	3,915,557,251,894	24.69%
Samsung	I9100 Galaxy S II	1,110,123,728,000	7.00%
Samsung	I337M	1,065,424,086,144	6.72%
Apple	iPhone 4	1,003,978,142,068	6.33%
Samsung	I9300 Galaxy S III	929,427,976,239	5.86%
Novatel Wireless	MiFi 2372	668,607,818,181	4.22%
Samsung	I747M	652,763,873,910	4.12%
HTC	One	629,293,358,808	3.97%
Samsung	I317M	311,537,719,257	1.96%
HTC	One X	283,633,356,106	1.79%
Apple	iPhone 3GS	201,431,941,002	1.27%
Apple	iPhone 5S	169,450,151,877	1.07%
Apple	iPhone 5C	132,042,611,201	0.83%
LG	P920h	127,230,100,578	0.80%
Novatel Wireless	Ovation MC547	106,706,160,962	0.67%
Samsung	I717 Galaxy Note	88,703,002,546	0.56%
HTC	PM36100	73,045,460,231	0.46%
Sony	LT28i Xperia Ion	65,916,567,474	0.42%
Samsung	I9000 Galaxy S	55,565,160,192	0.35%
Samsung	S5830L	48,478,046,405	0.31%
Samsung	I9000T	40,664,447,328	0.26%
SonyEricsson	R800a Xperia Play	40,446,228,545	0.26%
HTC	Windows Phone 8X	38,805,248,969	0.24%
Apple	iPad mini	33,233,414,392	0.21%

Units measured in bytes.

Figure 6 - Network Demographics report showing bytes attributable to specific mobile device models



Top Protocols per Access Device

Cluster: ...
Element: *
Policy Expression Instance(s): Apple|iPhone 5|*, Novatel Wireless|MiFi 2372|*
Date: 2013-10-02 00:00 -- 2013-10-09 00:00

Top 25 Manufacturer | Model | Protocol(s)

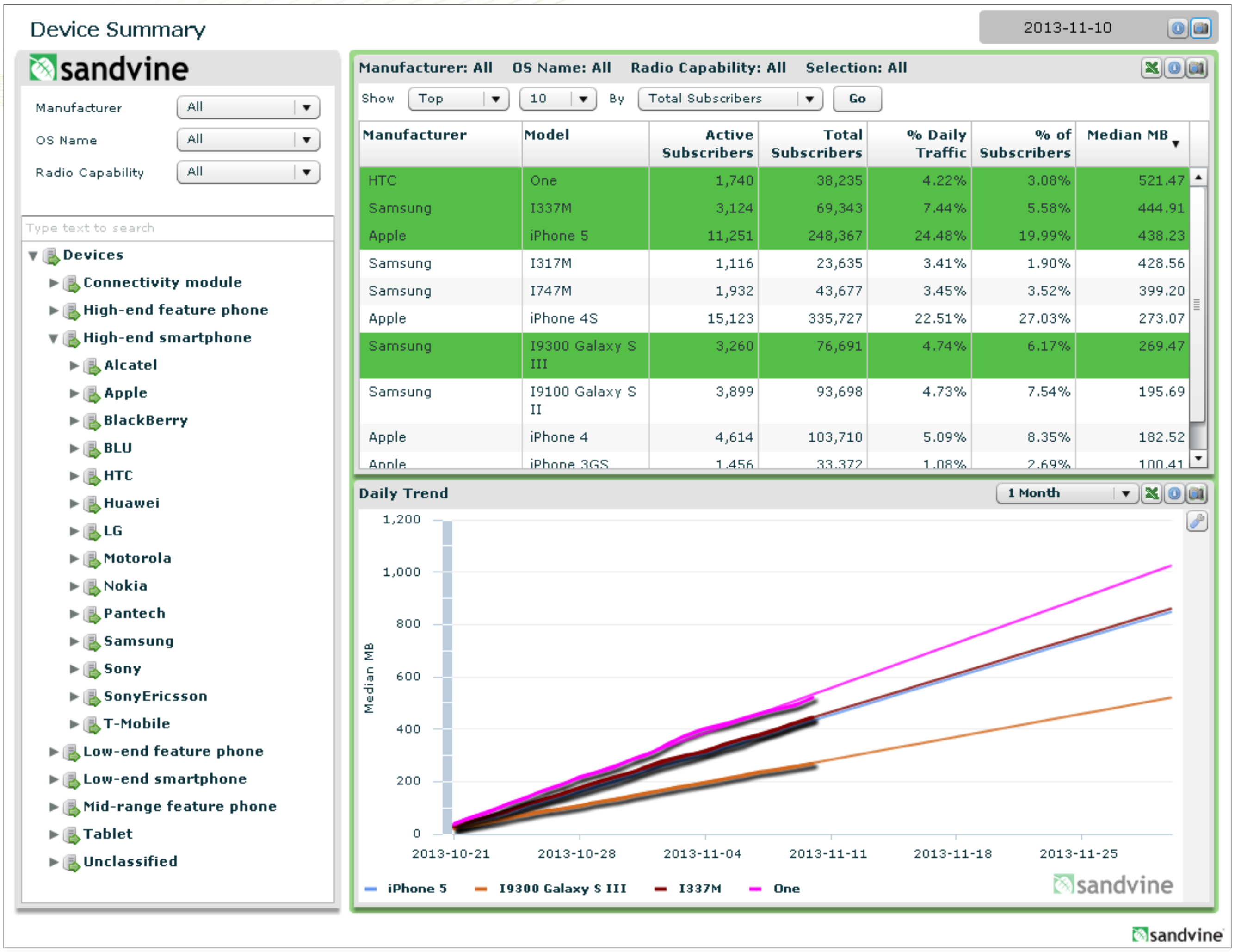
Manufacturer: Apple

Model: iPhone 5

Protocol	Total	Distribution
Facebook	621,969,880,392	16.90%
YouTube	558,852,505,225	15.18%
HTTP	553,608,734,624	15.04%
MPEG	261,889,988,706	7.12%
iTunes App Store	240,400,081,648	6.53%
Instagram	190,663,422,724	5.18%
SSL	174,708,266,871	4.75%
Netflix	173,099,415,438	4.70%
iTunes Streaming	151,399,552,712	4.11%
PANDORA Radio	136,333,847,294	3.70%
iTunes	111,802,020,683	3.04%
tumblr	66,551,390,584	1.81%
Apple Updates	60,939,441,027	1.66%
HTTP Live Streaming (HLS)	52,005,666,878	1.41%
Vine	47,585,685,765	1.29%
FaceTime	41,178,744,476	1.12%
Amazon	37,905,337,540	1.03%
Other TCP Protocol	36,210,707,808	0.98%
Skype PC to PC	30,839,917,981	0.84%
Pinterest	23,594,526,070	0.64%
OOYALA	22,856,950,504	0.62%
Apple iMessage	22,481,872,472	0.61%
Existing or Unmanaged Session	22,058,866,987	0.60%
iCloud	20,899,291,042	0.57%
Twitter	20,577,521,438	0.56%

Units measured in bytes.

Figure 7 - Network Demographics report showing top protocols and applications for the Apple iPhone 5



Device Summary

2013-11-10

sandvine

Manufacturer: All
OS Name: All
Radio Capability: All

Type text to search

- Devices
  - Connectivity module
  - High-end feature phone
  - High-end smartphone
    - Alcatel
    - Apple
    - BlackBerry
    - BLU
    - HTC
    - Huawei
    - LG
    - Motorola
    - Nokia
    - Pantech
    - Samsung
    - Sony
    - SonyEricsson
    - T-Mobile
  - Low-end feature phone
  - Low-end smartphone
  - Mid-range feature phone
  - Tablet
  - Unclassified

Manufacturer: All OS Name: All Radio Capability: All Selection: All

Show: Top ▼ 10 ▼ By Total Subscribers ▼ Go

Manufacturer	Model	Active Subscribers	Total Subscribers	% Daily Traffic	% of Subscribers	Median MB ▼
HTC	One	1,740	38,235	4.22%	3.08%	521.47
Samsung	I337M	3,124	69,343	7.44%	5.58%	444.91
Apple	iPhone 5	11,251	248,367	24.48%	19.99%	438.23
Samsung	I317M	1,116	23,635	3.41%	1.90%	428.56
Samsung	I747M	1,932	43,677	3.45%	3.52%	399.20
Apple	iPhone 4S	15,123	335,727	22.51%	27.03%	273.07
Samsung	I9300 Galaxy S III	3,260	76,691	4.74%	6.17%	269.47
Samsung	I9100 Galaxy S II	3,899	93,698	4.73%	7.54%	195.69
Apple	iPhone 4	4,614	103,710	5.09%	8.35%	182.52
Apple	iPhone 3GS	1,456	33,372	1.08%	2.69%	100.41

Daily Trend

1 Month ▼

Figure 8 - The Device Insights Dashboard within Network Analytics shows a wide range of statistics, measurements, and trends, broken down by a range of device characteristics

Conclusion

Sandvine’s traffic classification technology overcomes the challenges associated with device awareness to provide CSPs with the insight to understand the rich assortment of devices on their network and to empower CSPs with the policy control capabilities to profitably manage the Internet of Things.

Our traffic classification technology:

- Identifies individual client devices in real-time
- Provides detailed measurements and metrics per device
- Differentiates between client and access devices
- Detect device tethering and lets CSPs apply separate policy management to the tethered device versus the access device
- Identifies devices behind NATs
- Links all measured conditions to real-time policy control (i.e., decision-making and policy enforcement), enabling a vast array of powerful use cases

Headquarters
Sandvine Incorporated ULC
Waterloo, Ontario Canada
Phone: +1 519 880 2600
Email: sales@sandvine.com

European Offices
Sandvine Limited
Basingstoke, UK
Phone: +44 0 1256 698021
Email: sales@sandvine.co.uk

sandvine®
Intelligent Broadband Networks